NUMERICAL ESTIMATION OF SEVERAL TOPOLOGICAL QUANTITIES OF THE FIRST PASSAGE PERCOLATION MODEL

A Dissertation
Presented to
The Academic Faculty

By

Yuanzhe Ma

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Mathematics

Georgia Institute of Technology

May 2021

NUMERICAL ESTIMATION OF SEVERAL TOPOLOGICAL QUANTITIES OF THE FIRST PASSAGE PERCOLATION MODEL

Approved by:

Dr. Michael Damron, Advisor
School of Mathematics
Georgia Institute of Technology

Dr. Christian Houdré
School of Mathematics
Georgia Institute of Technology

Dr. Konstantin Tikhomirov
School of Mathematics
Georgia Institute of Technology

Date approved: April 5, 2021

ACKNOWLEDGMENTS

I would like to thank my advisor, Prof. Michael Damron, for introducing me to the research topic of first-passage-percolation. This work would not have been possible without Prof. Damron, his constant encouragement, wisdom, and expertise in probability always keep me highly motivated. I am always fascinated by probability and randomness, and this thesis project enables me to explore relevant theoretical papers along with doing interesting numerical experiments. I enjoy each of our conversations throughout the project, as I can always obtain new insights into our problem. I would also like to thank Prof. Damron for his excellent course in stochastic calculus, which is one of the best and clearest mathematics courses I have taken at Georgia Tech. In addition, I wish to thank Prof. Christian Houdré and Prof. Konstantin Tikhomirov for being on the thesis reading committee and providing much useful feedback.

This research was supported in part through research cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology, Atlanta, Georgia, USA.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In this thesis, our main goal is to use numerical simulations to study some quantities related to the growing set $B(t)$. Motivated by prior works, we mainly study quantities including the boundary size, the hole size, and the location of each hole for $B(t)$. We discuss the theoretical background of this work, the algorithm we used to conduct simulations, and include an extensive discussion of our simulation results. Our results support some of the prior conjectures and further introduce several interesting open problems.

CHAPTER 1
INTRODUCTION

We consider several quantities related to the growing set of first-passage-percolation (FPP), which was originally introduced by [1] and many recent results can be found in [2].

The model was originally used to model the fluid flow through a random material. It is also related to spread of an epidemic [3].

In addition, FPP is related to longest common subsequences (LCS) of two random words, which can be used to model DNA and protein sequences, a connection appears below the definition of V_n on page 1076 of [4]. More discussions of LCS and its relation to FPP can be found in [5, 6]. FPP can also be shown to be related to rumor spreading models [7]. One special FPP model, called the Eden model, which will be introduced later, can be used to model cell growth [8].

Despite its simple definition and great potential in applications, many important properties and statistics of this model are not well understood.

The organization of this thesis is as follows: chapter 2 introduces the basic definition and important theoretical background for this model, chapter 3 discusses the high-level procedure we are using and introduce six quantities which we are interested in, chapter 4 gives extensive remarks based on our simulation results, and finally chapter 5 gives conclusions and points out possible future directions.

CHAPTER 2

BACKGROUND

In this chapter, we will give the model definition and explain relevant notations and motivations for our work. We consider the graph defined on the lattice $(\mathbb{Z}^d, \mathcal{E}^d)$ for $d \geq 2$, where $\mathcal{E}^d$ is the collections of nearest-neighbor edge in $\mathbb{Z}^d$. We consider a sequence of independent, identically distributed (i.i.d.) non-negative random variables t_e and we assign t_e to edge $e \in \mathcal{E}^d$. We also denote the common distribution of $(t_e)_{e \in \mathcal{E}^d}$ as F. We then define the path using a sequence of vertices and edges $(x_0, e_1, x_1, \cdots, e_n, x_n)$, here $x_i \in \mathbb{Z}^d$ are vertices and $e_i \in \mathcal{E}^d$ is the edge connecting x_{i-1} and x_i. Given two vertices $x, y \in \mathbb{Z}^d$, the first-passage-time from of x to y is then defined as

$$T(x, y) = \inf_{\gamma: x \to y} T(\gamma),$$

here γ is any path starting from x to y and $T(\gamma) = \sum_{e \in \gamma} t_e$ which can be considered the total time or weight to travel from x to y following path γ.

The growing set $B(t)$ can be defined as

$$B(t) = \{x \in \mathbb{Z}^d : T(0, x) \leq t\}.$$

2.1 Theoretical results

In this section, we present several important theoretical results.

Theorem 1 (Theorem 2.1 of [2]) *Suppose* $\mathbb{E} \min\{t_1, \cdots, t_{2d}\} < \infty$, *where* $t_i, i = 1, \cdots, 2d$ *are independent copies of* t_e, *then there exists a constant (called the time constant)* $\mu(e_1) \in [0, \infty)$ *such that*

$$\lim_{n \to \infty} \frac{T(0, ne_1)}{n} = \mu(e_1) \text{ a.s. and in } L^1.$$

The theorem characterizes the limiting behavior of $\frac{T(0,ne_1)}{n}$ and the proof is based on subadditive ergodic theorem, more detail can be seen in [2].

Another important result of $B(t)$ is the shape theorem, which characterizes the limiting behavior of $\frac{B(t)}{t}$. Before presenting this result, we need to define $T(x,y)$ for $x, y \in \mathbb{R}^d$. Given $x \in \mathbb{R}^d$, we let $[x]$ be the unique element in $\mathbb{Z}^d$ such that $x \in [x] + [0,1)^d$, then we can define $T(x,y) = T([x],[y])$ given $x, y \in \mathbb{R}^d$. In other words, we just extended the domain of T from $\mathbb{Z}^d \times \mathbb{Z}^d$ to $\mathbb{R}^d \times \mathbb{R}^d$.

Theorem 2 (Theorem 2.16 of [2]) *Suppose* $\mathbb{E}\min\{t_1^d, \cdots, t_{2d}^d\} < \infty$ *and* $F(0) < p_c(d)$ *where* $p_c(d)$ *is the bond percolation threshold for* d *dimensions. Then there exists a compact, convex, and deterministic set* $\mathcal{B}$ *in* $\mathbb{R}^d$ *such that for every* $\epsilon > 0$,

$$\mathbb{P}\left((1-\epsilon)\mathcal{B} \subset \frac{\bar{B}(t)}{t} \subset (1+\epsilon)\mathcal{B} \text{ for all large } t\right) = 1,$$

where $\bar{B}(t) = \{x \in \mathbb{R}^d : T(0,x) \leq t\}$. *In addition,* $\mathcal{B}$ *has non-empty interior and is symmetric about the axes of* $\mathbb{R}^d$. *Moreover, the conditions* $\mathbb{E}\min\{t_1^d, \cdots, t_{2d}^d\} < \infty$ *and* $F(0) < p_c(d)$ *are necessary for the above equality to hold.*

We now introduce two types of boundaries which will be used later.

Definition 1 (Definition 1.1 of [9]) *Given a graph* $(V, \mathcal{E}^d)$ *with* $V \subset \mathbb{Z}^d$.

1. *We define the edge boundary of* V *as*

$$\#\partial_e V = \{\{x,y\} \in \mathcal{E}^d : x \in V, y \in \mathbb{Z}^d \setminus V\}$$

2. *We define the exterior boundary of* V, $\partial^{\text{ext}} V$ *of* $V \subset \mathbb{Z}^d$ *as the set of all* $x \in \mathbb{Z}^d \setminus V$ *that are*

   (a) *adjacent to a vertex in* V, *and* *(b) the starting point of some infinite vertex self-avoiding path which does not intersect V.*

The edge exterior boundary of $\partial_e^{\text{ext}} V$ of a set $V \subset \mathbb{Z}^d$ is the set of edges $\{x, y\}$ for some $y \in V$ and $x \in \partial^{\text{ext}} V$.

A natural question is what is the order of $\#\partial_e B(t)$. It is conjectured that $\#\partial_e B(t) \sim t^{d-1}$ [9]. However, there are no rigorous results proving that t^{d-1} is the correct order of $\#\partial_e B(t)$ for all large t.

And it turns out this is related to a random variable $Y = \min\{t_1, \cdots, t_{2d}\}$ and it was shown in [9] that if $\mathbb{E}Y < \infty$ then for most times t one almost surely has $\#\partial_e B(t) \sim t^{d-1}$.

More precisely, we have the following results:

Theorem 3 (Theorem 1.2 of [9]) *Suppose that $\mathbb{P}(t_e = 0) < p_c(d)$.*

Define

$$R_t(a) = \{s \in [0, t] : \#\partial_e B(s) \geq a s^{d-1} \mathbb{E}[Y \wedge s]\},$$

and

$$R_t^{\text{ext}}(a) = \{s \in [0, t] : \#\partial_e^{\text{ext}} B(s) \geq a s^{d-1}\}.$$

(a) There exists $C > 0$ such that almost surely,

$$\limsup_{t\to\infty} \frac{\text{Leb}(R_t(a))}{t} \leq \frac{C}{a}.$$

(b) There exists $C > 0$ such that almost surely,

$$\limsup_{t\to\infty} \frac{\text{Leb}(R_t^{\text{ext}}(a))}{t} \leq \frac{C}{a}.$$

Several direct consequences of this theorem [9] are:

1. If $\mathbb{E}[Y] < \infty$, then Theorem 3 (a) implies that $\#\partial_e B(t) \leq a\mathbb{E}[Y]t^{d-1}$ for most t.

2. If $\mathbb{E}[Y] = \infty$ and there exists a constant $C > 0$ such that $\mathbb{P}\left(Y \geq y\right) \leq \frac{C}{y}$ for every $y > 0$, then Theorem 3 (a) implies that $\#\partial_e B(t) \leq at^{d-1}\log t$ for most t.

3. If there exists a constant $C > 0$ such that $\mathbb{P}\left(Y \geq y\right) \leq \frac{C}{y^{1-\alpha}}$ for every $y > 0$ and some $\alpha \in (0,1)$, then Theorem 3 (a) implies that $\#\partial_e B(t) \leq at^{d-1+\alpha}$ for most t.

There is also one result for the lower bound of $\#\partial_e B(t)$:

Theorem 4 (Theorem 1.3 of [9]) *Suppose that $\mathbb{P}\left(t_e = 0\right) < p_c(d)$ and let F_Y be the distribution function of Y. There exists $C > 0$ such that almost surely,*

$$\#\partial_e B(t) \geq C\left[(1 - F_Y(t)) \vee \frac{1}{t}\right] t^d \text{ for all large } t.$$

Combining the above two results, we can say that given $\mathbb{P}\left(t_e = 0\right) < p_c(d)$, for most values of t,

$$[t(1 - F_Y(t)) \vee 1]t^{d-1} \lesssim \#\partial_e B(t) \lesssim \mathbb{E}[Y \wedge t]t^{d-1}.$$

Theorem 3 implies that the size of the exterior boundary of $B(t)$ is usually at most of order t^{d-1}. However, Theorem 4 means that for some heavy-tail distributions of t_e, the full size of the boundary is much larger under low moments. Thus, we conclude there must be many holes that contribute to the boundary size. One [9] can also show that most of these holes must be small, so there should be many small holes inside $B(t)$. Below we give a formal definition of a hole for a graph defined on $\mathbb{Z}^d$.

Definition 2 *Consider a graph $G = (V, E)$ defined on $(\mathbb{Z}^d, \mathcal{E}^d)$. Two vertices u and v of G are called connected if and only if there is a path connecting u and v. We also define a component of G as a maximal set all of whose pairs of vertices are connected. Let G^C be the complement of G. A hole of G is a component of G^C with finite vertices.*

However, the previous results do not directly imply an almost sure result of the order of $\#\partial_e B(t)$ for large t. We wish to use the simulation results to verify this, which is one of the open questions in Section 4 of [9]. Understanding the behaviors of holes for each distribution is also one purpose of the simulation.

Suppose $\mathcal{B}$ satisfies a condition called uniform curvature condition [9, Definition 1.4], $\mathbb{P}(t_e = 0) < p_c(d)$, and $\mathbb{E}[e^{\alpha t_e}] < \infty$ for some $\alpha > 0$, then one [9, Theorem 1.5] can show that $\#\partial_e B(t) \leq C(\log t)^C t^{d-1}$ for some $C > 0$ almost surely for all large t. The uniform curvature is believed to be true for t_e having continuous distributions, but no formal proof exists, see more discussions in [2, Section 2.8].

We are also interested in the location of each hole, and we believe most holes should be close to the boundary for large enough t with high probability for certain weight distributions. Thus, we would also like to understand the distribution of the distance between holes and exterior boundary points. Here for two sets A and B, we define $\mathrm{dist}(A, B) = \inf\{\|x - y\|_2, x \in A, y \in B\}$.

More recent theoretical results related to the number of holes of $B(t)$ can be seen at Theorem 1 of [10].

2.2 Prior results of simulations

The Eden model is a growth model introduced by [8] and was shown to be equivalent to FPP with t_e following exponential distribution [11]. Simulation results for the time constants of Eden model and other FPP models with general weight distributions in $\mathbb{Z}^2$ are provided in [12].

More recent simulation results for various topological quantities of the Eden model can be seen in [10]. However, no simulation results related to topological quantities related to general FPP are known.

CHAPTER 3
SIMULATION PROCEDURE

In this section, we discuss our procedure to conduct the simulations, due to the constraint of computational power, we mainly focus on the simulations of $\mathbb{Z}^2$. However, our procedure could be extended to $\mathbb{Z}^d$.

We summarize the quantities that we are interested in and their abbreviations in this thesis:

1. the size of boundary of $B(t)$: boundary_size
2. the number of holes inside $B(t)$: n_hole
3. the total volume of holes inside $B(t)$: total_hole_size
4. the largest volume of holes inside $B(t)$: max_hole_size
5. the average ℓ_2 distance between each hole of $B(t)$ and the exterior boundary of $B(t)$: avg_hole_boundary_distance
6. the largest ℓ_2 distance between each hole of $B(t)$ and the exterior boundary of $B(t)$: max_hole_boundary_distance

Note that all quantities above are continuous-time stochastic processes.

For each of the quantity defined above, we will do 4000 simulations and would like to investigate the relationship between its mean and standard deviation versus time t. Note that by the definition of $T(x, y)$, for general distributions, it is impossible to compute $T(x, y)$ without errors. Thus, when conducting simulations, we first create a large enough box, and only define edge weights inside the large box.

The pseudo algorithm for one simulation can be described as follows:

Algorithm 1: Compute all quantities with edge weight distribution $F(\cdot)$

Initialize a lattice graph G with vertices $[-1500, 1500]^2$ in $\mathbb{Z}^2$;

For each edge e in G, create a random variable $t_e \sim F(\cdot)$;

Use Dijkstra's shortest path algorithm to compute $T(0, x)$;

for *each* t **do**

    Define the set $B(t) = \{x \in \mathbb{Z}^2 : T(0, x) \leq t\}$;

    Find all holes in $B(t)$, which are finite components of $B(t)^C$;

    For each vertex in $B(t)$, compute its degree, those with degree less or equal to 3 are marked as boundary points, note that some points in holes can also be boundary points;

    For each vertex in $B(t)^C$, compute its degree, those with degree less or equal to 3 and are not part of holes in $B(t)$ are marked as exterior boundary points

end

CHAPTER 4

RESULTS

In this chapter we present the results of the simulations.

We mainly consider the quantities defined in the previous chapter for three different distributions for the edge weight distributions t_e:

1. Exponential distribution with mean 1, and we use Exp (1) to denote this distribution

2. Pareto distribution with parameter 0.5, i.e. the probability density function $f(x) \propto \frac{1}{x^{1.5}}$ for $x \geq c$ where c is a constant, and we use Pareto (0.5) to denote this distribution

3. Pareto distribution with parameter 0.1, i.e. the probability density function $f(x) \propto \frac{1}{x^{1.1}}$ for $x \geq c$ where c is a constant, and we use Pareto (0.1) to denote this distribution

We now present several sample plots of $B(t)$ for these distributions. In each plot, we use the blue region to indicate $B(t)$, red dots to represent its holes, the green line to represent the boundary of $B(t)$, and the yellow line to stand for its exterior boundary.

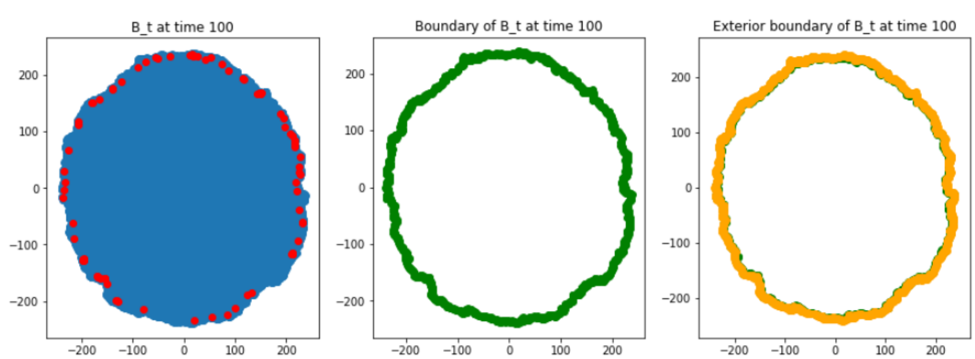


Figure 4.1: Sample plot of $B(t)$ for Exp (1)

For each distribution, we will run 4000 simulations, and compute 6 quantities defined in chapter 3 on a given time interval with some step size. Then for each quantity we estimate

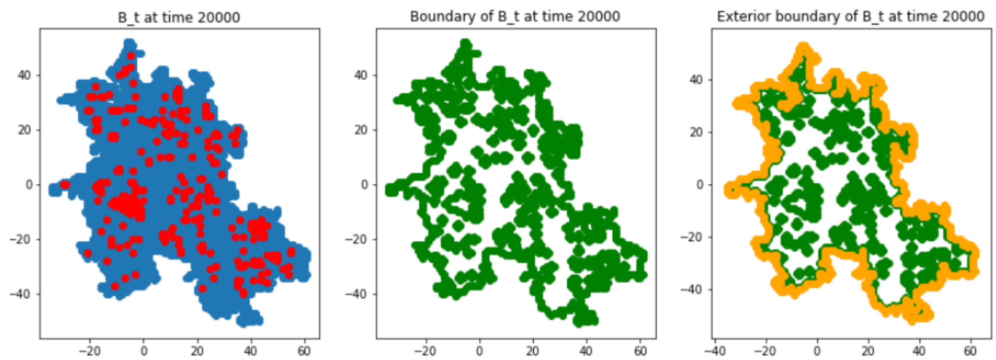


Figure 4.2: Sample plot of $B(t)$ for Pareto (0.1)

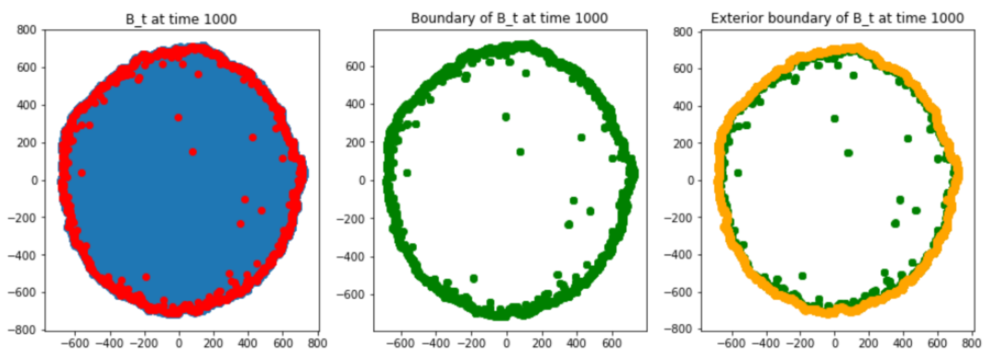


Figure 4.3: Sample plot of $B(t)$ for Pareto (0.5)

its mean and standard deviation for 4000 samples. Finally, we use least square regression to approximate it in the form of t^α or $(\log t)^\alpha$.

More precisely, for each variable $x(t)$, we estimate its mean $\bar{x}(t)$, we fit the parameter using models in the form of

$$\hat{\alpha}, \hat{a}, \hat{b}, \hat{c} = \underset{\alpha,a,b,c\in\mathbb{R}}{\arg\min} \sum_t \left(\bar{x}(t) - at^\alpha - bt^{\alpha-1} - c\right)^2,$$

or

$$\hat{\alpha}, \hat{a}, \hat{b}, \hat{c} = \underset{\alpha,a,b,c\in\mathbb{R}}{\arg\min} \sum_t \left(\bar{x}(t) - a(\log t)^\alpha - b(\log t)^{\alpha-1} - c\right)^2,$$

here we will choose to use t or $\log t$ based on prior knowledge and the model results.

We include an extra $\alpha - 1$ term for numerical purposes, and it turns out this does not influence the results much.

For each variable, we obtain the power estimation $\hat{\alpha}$ using the above equation and present the results in the following two tables. In addition, we put all results in plots in the appendix.

Table 4.1: Simulation results for the mean of each variable

Variable	Exp (1)	Pareto (0.5)	Pareto (0.1)
boundary_size	$t^{1.009}$	$t^{1.004}$	$t^{1.690}$
n_hole	$t^{1.000}$	$t^{1.006}$	$t^{1.819}$
total_hole_size	$t^{1.028}$	$t^{1.003}$	$t^{1.571}$
max_hole_size	$(\log t)^{1.988}$	$(\log t)^{1.248}$	$(\log t)^{1.003}$
avg_hole_boundary_distance	$(\log t)^{0.582}$	$(\log t)^{1.846}$	$t^{1.000}$
max_hole_boundary_distance	$(\log t)^{0.023}$	$t^{1.016}$	$t^{1.099}$

Table 4.2: Simulation results for the standard deviation of each variable

Variable	Exp (1)	Pareto (0.5)	Pareto (0.1)
boundary_size	$t^{0.438}$	$t^{0.521}$	$t^{0.734}$
n_hole	$t^{0.287}$	$t^{0.997}$	$t^{1.000}$
total_hole_size	$t^{0.998}$	$t^{0.980}$	$t^{0.559}$
max_hole_size	$(\log t)^{0.479}$	$(\log t)^{0.994}$	$(\log t)^{1.006}$
avg_hole_boundary_distance		$(\log t)^{0.460}$	$(\log t)^{0.676}$
max_hole_boundary_distance	C	$t^{1.000}$	$(\log t)^{1.000}$

Here we give several remarks to explain the above results:

1. In the first table, the entry $t^{1.009}$ indicates we believe for Exp (1), the mean of boundary size at time t can be approximated as $at^{1.009} + bt^{0.009} + c$ for some constants a, b, c. Other entries can be understood accordingly.

2. We believe the standard deviation of avg_hole_boundary_distance at time t is a decreasing function of t for Exp (1) but there is no reasonable approximation function for this variable based on the data we collect. Therefore, we leave this entry blank.

3. We believe the standard deviation of max_hole_boundary_distance at time t is nearly a constant for Exp (1), if t is sufficiently large.

4. For some variables, we are restricting $a, b \geq 0$ to improve numerical stability.

5. We believe all powers in the range $[-0.95, 1.05]$ can be interpreted as 1. And we believe that we do not obtain exactly 1 because we cannot collect infinite amount of data points and there is inevitable small errors caused by our simulations, which is explained in chapter 3.

6. The condition for the shape theorem (Theorem 2) to hold is $\int_0^\infty x^{1-4\alpha} \mathrm{d}x < \infty$ for Pareto distribution with parameter α. In other words, we need $\alpha > 0.5$ for the shape theorem to hold. Therefore, for Pareto (0.5), we should expect some holes far from the boundary, and there should be more holes of this kind for Pareto (0.1). This is also confirmed in the results for avg_hole_boundary_distance.

7. Based on the results of the mean, for Pareto (0.1), the order of n_hole is larger than boundary_size. However, this is due to numerical issues. In fact, the coefficient a of boundary_size is much larger than that of n_hole. This means the quantity of boundary_size is still larger than n_hole. And we expect the order of n_hole should be smaller if we can simulate up to an extremely large time t, but this is not possible as of now due to computational constraints.

8. For Exp (1), all terms with estimated powers in log are quite small in scale and as a result, the results do not seem plausible. In fact, based on the plots in the appendix, it appears that all these terms should be close to $\log t$. However, since these quantities are so small, and we cannot verify this easily.

9. The simulation results support the conjecture that $\#\partial_e B(t) \sim t^{d-1}$ for certain distributions in [9].

10. For boundary_size, n_hole, and total_hole_size, the results of the mean seem to be similar for three different distributions, except for the estimated powers. This holds both for the mean and the standard deviation. We can also see that Pareto (0.1) seems to yield larger boundary size, more and larger holes, which are expected.

11. For the other variables including avg_hole_boundary_distance and max_hole_boundary_distance, the results of the mean are quite different for different distributions. There is clearly a transition from $\log t$ to t. For max_hole_size, the order seems to be decreasing from Exp (1) to Pareto (0.1), but as we can see in Figure 4, this is due to numerical issues and in magnitude, Pareto (0.1) has the largest max_hole_size generally. We believe these indicate with Pareto (0.1), or in general, heavy-tail distributions for the weight, the maximum hole size will be much larger, and there are many holes which are far way from the boundary.

12. For standard deviation results of the last three variables, we do not have a clear explanation yet, and this can be interesting future research directions. In addition, one should note that for max_hole_boundary_distance, the standard deviation is larger for Pareto (0.5), compared to Pareto (0.1), which is quite surprising.

CHAPTER 5
CONCLUSIONS

In this thesis, we investigate the properties of the random set $B(t)$ generated by different distributions. Relevant simulations support some conjectures related to the boundary size of $B(t)$.

Interesting future directions could be:

1. prove these results of the boundary size rigorously
2. try to understand the mean and the standard deviation of all other quantities
3. extend the analysis to the model with higher dimensions

It is clear that for simulations result to be more convincing, one should do the following:

1. conduct more simulations for each variable
2. increase the size of the ”box” we introduced in Algorithm 1
3. increase the time interval for the simulations

In this thesis, we only considered six variables, and it is worthy to consider more variables to understand the growth set $B(t)$ better. In addition, one can also do simulations for a class of weight distributions and try to understand the relation of each variable and the weight distribution. Here we only completed simulations for three different distributions, and their main difference is moments. However, maybe other distributions are worthy looking at. One recent work [13] is trying to understand the limiting shape for different weight distributions using neural networks. We think this can also be done for other quantities.

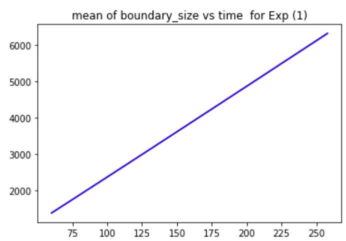
Besides the above suggestions, we would be happy to hear more ideas.

Appendices

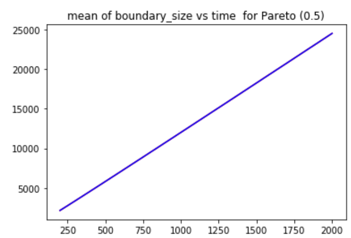
APPENDIX A

PLOTS OF RESULTS

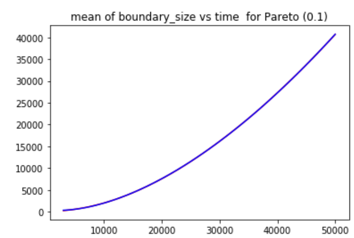
In this section, we present all the results in plots with their fitted powers. In each plot, the red line means the result of our estimation and the blue line represents our fitted polynomial in t or $\log t$.



(a) Exp (1): $t^{1.009}$

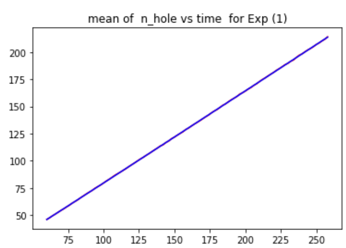


(b) Pareto (0.5): $t^{1.004}$

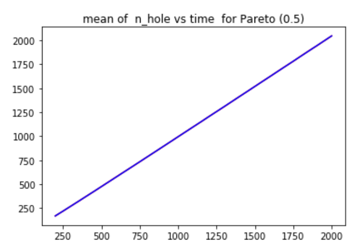


(c) Pareto (0.1): $t^{1.690}$

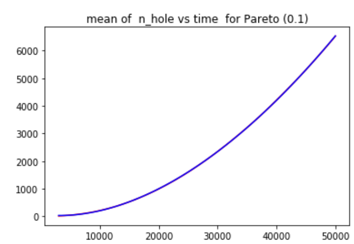
Figure A.1: Mean of boundary_size



(a) Exp (1): $t^{1.000}$

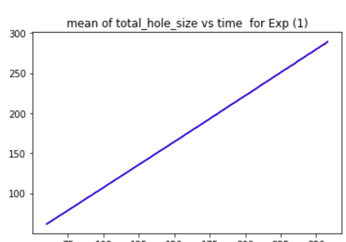


(b) Pareto (0.5): $t^{1.006}$

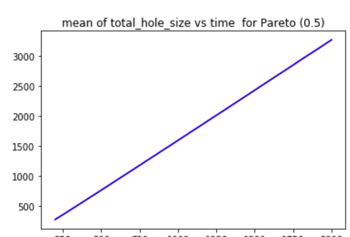


(c) Pareto (0.1): $t^{1.819}$

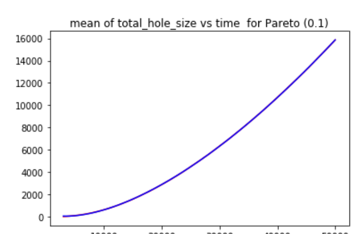
Figure A.2: Mean of n_hole



(a) Exp (1): $t^{1.028}$

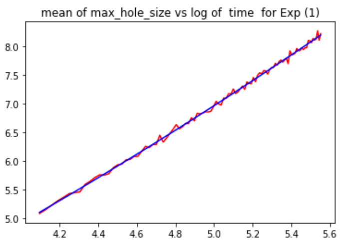


(b) Pareto (0.5): $t^{1.003}$

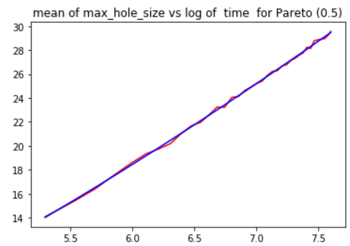


(c) Pareto (0.1): $t^{1.571}$

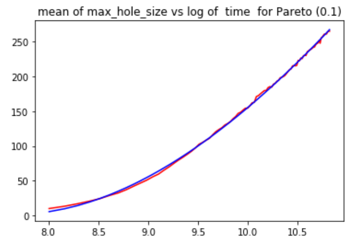
Figure A.3: Mean of total_hole_size



(a) Exp (1): $(\log t)^{1.988}$

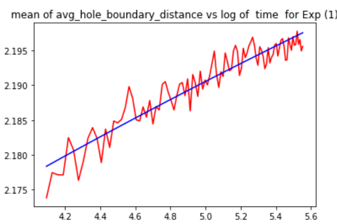


(b) Pareto (0.5): $(\log t)^{1.248}$

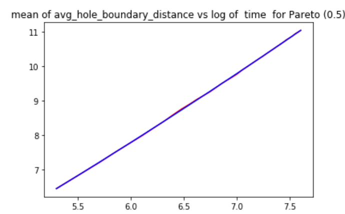


(c) Pareto (0.1): $(\log t)^{1.003}$

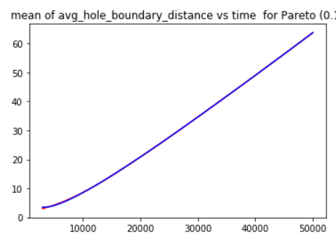
Figure A.4: Mean of max_hole_size



(a) Exp (1): $(\log t)^{0.582}$

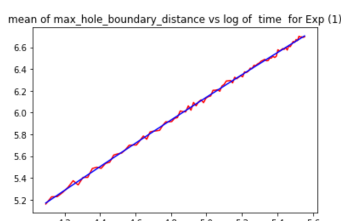


(b) Pareto (0.5): $(\log t)^{1.846}$

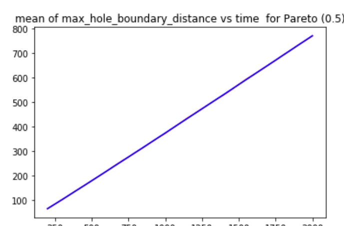


(c) Pareto (0.1):$t^{1.000}$

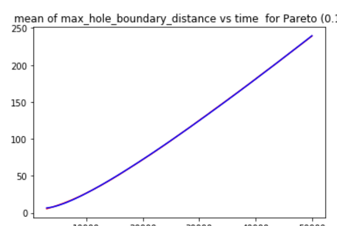
Figure A.5: Mean of avg_hole_boundary_distance



(a) Exp (1): $(\log t)^{0.023}$

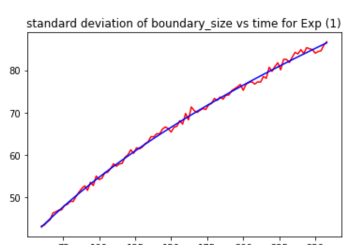


(b) Pareto (0.5): $t^{1.016}$

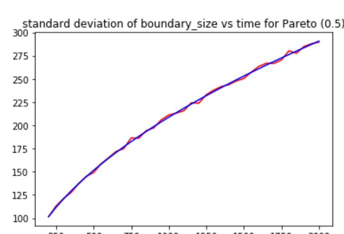


(c) Pareto (0.1): $t^{1.099}$

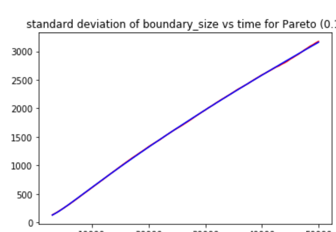
Figure A.6: Mean of max_hole_boundary_distance



(a) Exp (1): $t^{0.438}$

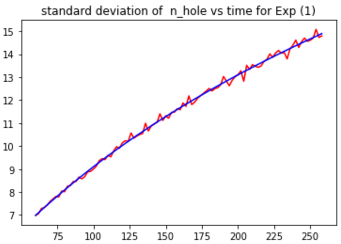


(b) Pareto (0.5): $t^{0.521}$

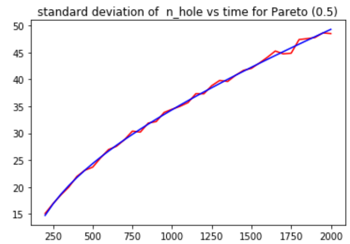


(c) Pareto (0.1): $t^{0.734}$

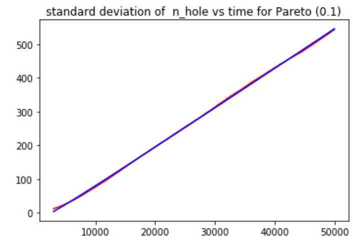
Figure A.7: Standard deviation of boundary_size



(a) Exp (1): $t^{0.287}$

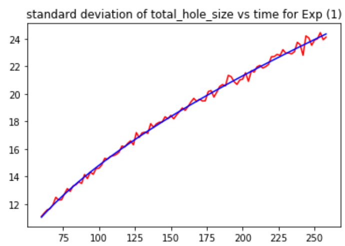


(b) Pareto (0.5): $t^{0.997}$

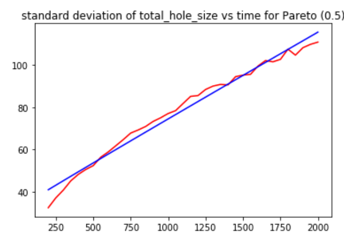


(c) Pareto (0.1): $t^{1.000}$

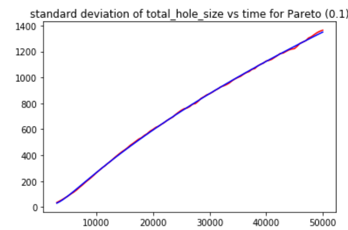
Figure A.8: Standard deviation of n_hole



(a) Exp (1): $t^{0.998}$

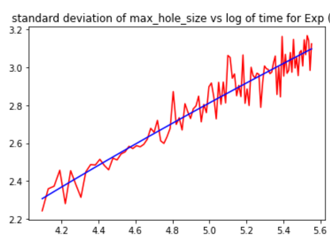


(b) Pareto (0.5): $t^{0.980}$

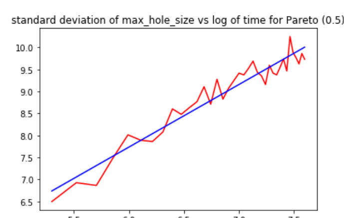


(c) Pareto (0.1): $t^{0.559}$

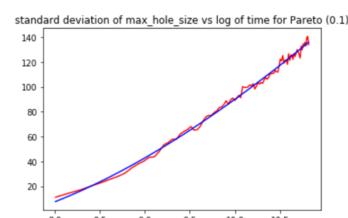
Figure A.9: Standard deviation of total_hole_size



(a) Exp (1): $(\log t)^{0.479}$

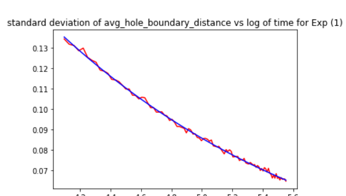


(b) Pareto (0.5): $(\log t)^{0.994}$

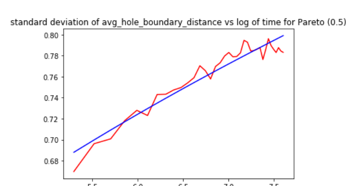


(c) Pareto (0.1): $(\log t)^{1.006}$

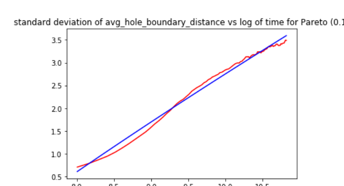
Figure A.10: Standard deviation of max_hole_size



(a) Exp (1)

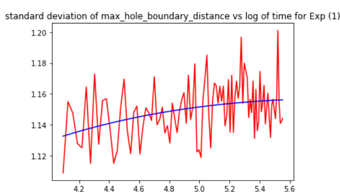


(b) Pareto (0.5): $(\log t)^{0.460}$

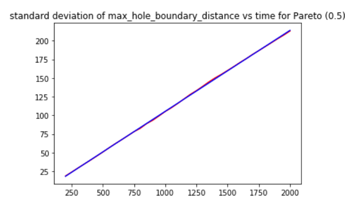


(c) Pareto (0.1): $(\log t)^{0.676}$

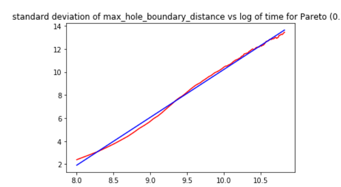
Figure A.11: Standard deviation of avg_hole_boundary_distance



(a) Exp (1): C



(b) Pareto (0.5): $t^{1.000}$



(c) Pareto (0.1): $(\log t)^{1.000}$

Figure A.12: Standard deviation of max_hole_boundary_distance

REFERENCES

[1] J. M. Hammersley and D. J. A. Welsh, "First-Passage Percolation, Subadditive Processes, Stochastic Networks, and Generalized Renewal Theory," in *Bernoulli 1713 Bayes 1763 Laplace 1813*, 1965.

[2] A. Auffinger, M. Damron, and J. Hanson, *50 years of first passage percolation*, 2016. arXiv: 1511.03262 `[math.PR]`.

[3] S. Bhamidi, R. Van Der Hofstad, and J. Komjáthy, "The front of the epidemic spread and first passage percolation," *Journal of Applied Probability*, vol. 51A, 2014.

[4] K. S. Alexander, "The Rate of Convergence of the Mean Length of the Longest Common Subsequence," *The Annals of Applied Probability*, vol. 4, no. 4, 1994.

[5] R. Arratia and M. S. Waterman, "A Phase Transition for the Score in Matching Random Sequences Allowing Deletions," *The Annals of Applied Probability*, vol. 4, no. 1, 1994.

[6] C. Houdré and H. Matzinger, "Closeness to the diagonal for longest common subsequences in random words," *Electronic Communications in Probability*, vol. 21, 2016.

[7] O. Angel, A. Mehrabian, and Y. Peres, "The string of diamonds is nearly tight for rumour spreading," *Combinatorics, Probability and Computing*, vol. 29, no. 2, pp. 190–199, 2020.

[8] M. Eden, "A two-dimensional growth process," *Berkeley Symposium on Mathematical Statistics and Probability*, vol. 4, 1961.

[9] M. Damron, J. Hanson, and W. K. Lam, "The size of the boundary in first-passage percolation," *Annals of Applied Probability*, vol. 28, no. 5, 2018.

[10] F. Manin, E. Roldan, and B. Schweinhart, *Topology and local geometry of the eden model*, 2020. arXiv: 2005.12349 `[math.PR]`.

[11] D. Richardson, "Random growth in a tessellation," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 74, no. 3, 1973.

[12] S. E. Alm and M. Deijfen, "First Passage Percolation on $\mathbb{Z}^2$: A Simulation Study," *Journal of Statistical Physics*, vol. 161, no. 3, pp. 657–678, Nov. 2015.

[13] S. Rosengren, *Predicting first passage percolation shapes using neural networks*, 2020. arXiv: 2006.14004 `[stat.ML]`.